Generable

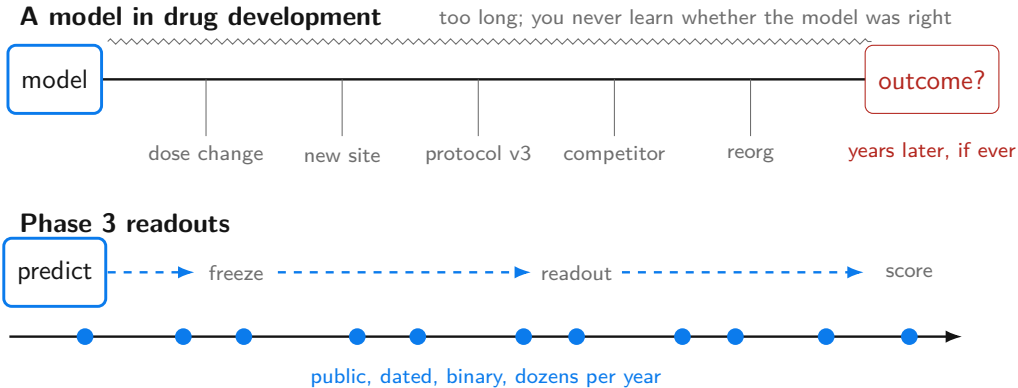
AI-assisted evidence synthesis with Bayesian meta-analysis for prospective trial forecasting

The ACACIA-HCM case study

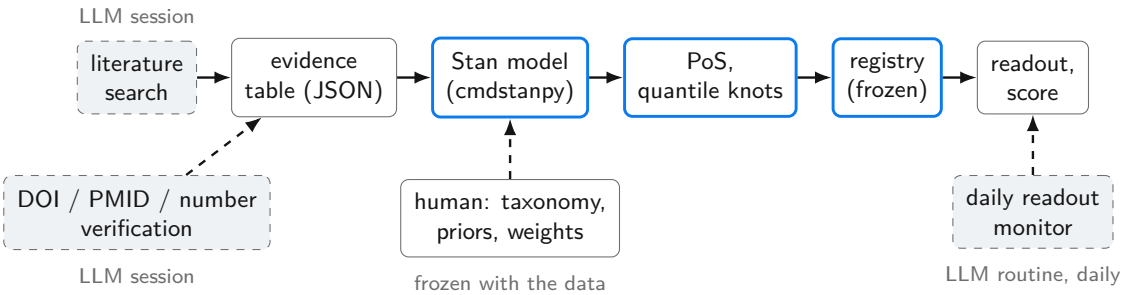
Eric Novik · Generable · NYU

Bayesian Biostatistics 2026, Leiden | October 2026

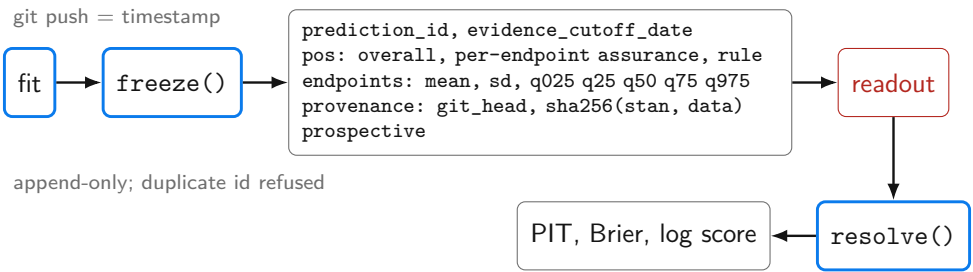
Why predict Phase 3 readouts?



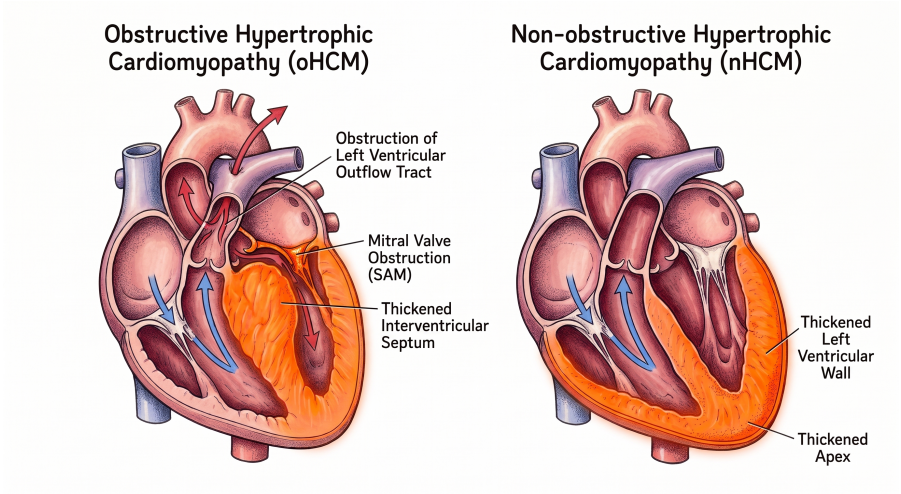
The workflow, and where the LLM sits



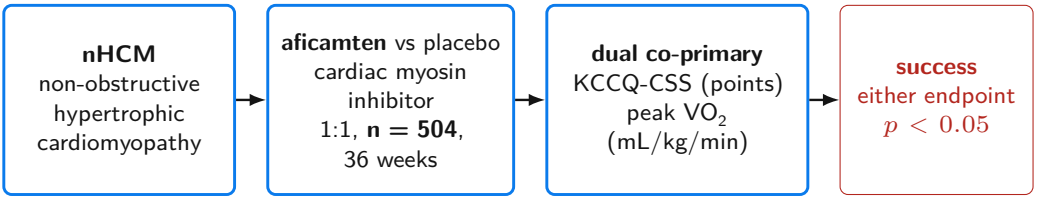
Freeze before readout: the registry



Hypertrophic Cardiomyopathy (HCM)

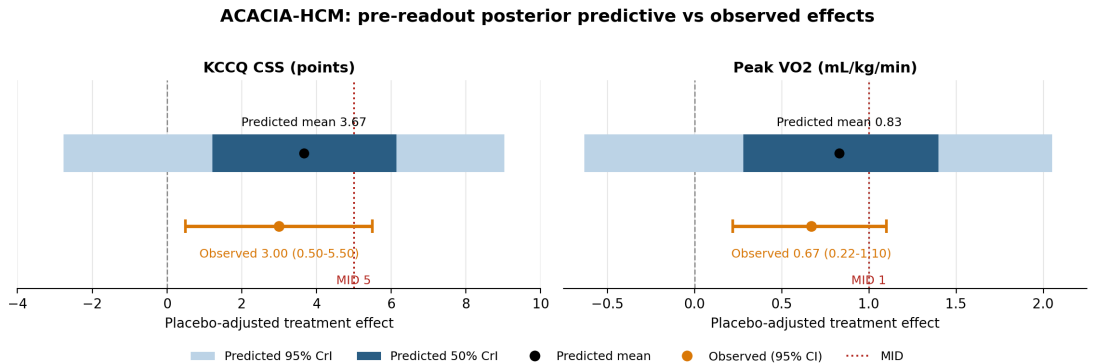


ACACIA-HCM: the trial

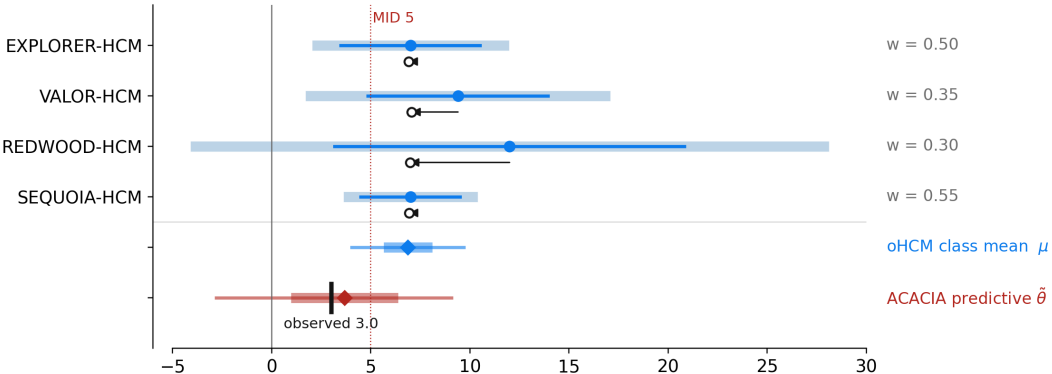


The transfer question: the same mechanism worked three times in *obstructive* HCM. Does it carry to the non-obstructive population, and by how much less?

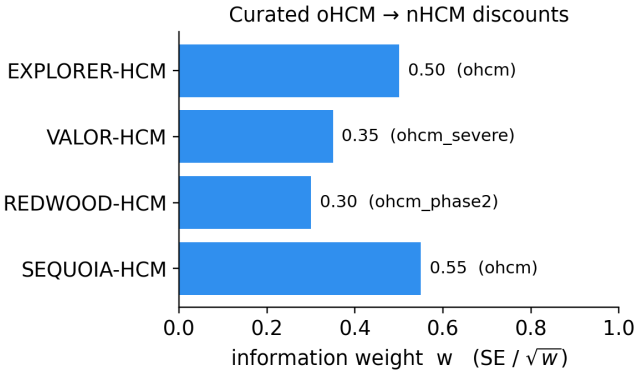
ACACIA-HCM: frozen April 22, read out May 5

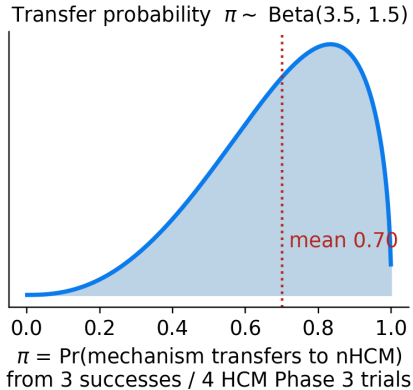


Prior evidence: four oHCM trials



Two prior ingredients





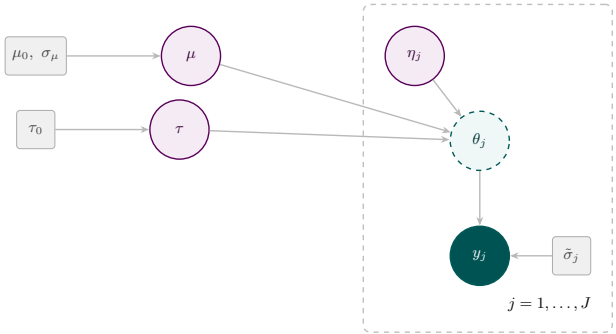
Model part 1: pooling the oHCM trials

$$y_j \mid \theta_j \sim \mathcal{N}\left(\theta_j,\, \tilde{\sigma}_j^2\right), \quad \tilde{\sigma}_j = \frac{\sigma_j}{\sqrt{w_j}}$$

$$\theta_j \mid \mu, \tau \sim \mathcal{N}(\mu, \tau^2)$$

$$\mu \sim \mathcal{N}(0, 5^2), \quad \tau \sim \text{HalfNormal}(1.5)$$

The discount w_j widens a trial's likelihood; it never moves its estimate.



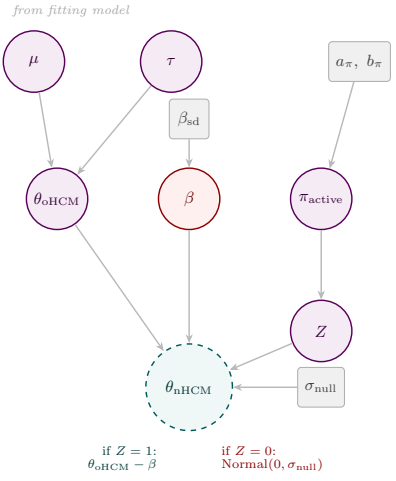
Model part 2: the nHCM predictive

$$\pi \sim \text{Beta}(3.5,\, 1.5), \qquad Z \sim \text{Bernoulli}(\pi)$$

$$\beta \sim \text{HalfNormal}(2), \qquad \tilde{\theta}_{\text{oHCM}} \sim \mathcal{N}(\mu, \tau^2)$$

$$\tilde{\theta}_{\text{nHCM}} = \begin{cases} \tilde{\theta}_{\text{oHCM}} - \beta & Z = 1 \\ \mathcal{N}(0, 2^2) & Z = 0 \end{cases}$$

π, Z, β have no likelihood term: the transfer is a prior.



Success rule: dual co-primary, either hits

single endpoint: $\Pr(\text{sig} \mid \tilde{\theta}) = \Phi\left(\dfrac{\tilde{\theta}}{\text{se}} - z_{0.975}\right), \qquad \text{se} = \sigma_w \sqrt{\tfrac{1}{252} + \tfrac{1}{252}}$

either: $\begin{pmatrix} \hat{\delta}_{\text{KCCQ}} \\ \hat{\delta}_{\text{VO}_2} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \tilde{\theta}_{\text{KCCQ}} \\ \tilde{\theta}_{\text{VO}_2} \end{pmatrix}, \begin{pmatrix} \text{se}_1^2 & \rho\,\text{se}_1\text{se}_2 \\ \rho\,\text{se}_1\text{se}_2 & \text{se}_2^2 \end{pmatrix} \right), \quad \rho = 0.37$

$$\text{PoS} = \mathbb{E}\left[\, \mathbb{1}\{z_{\text{KCCQ}} > 1.96 \text{ or } z_{\text{VO}_2} > 1.96\} \,\right]$$

Bayesian predictive power: the significance probability averaged over the posterior of the true effect. ρ is the correlation of the two endpoints' *sampling errors*.

Stan: data block

```
data {
  int<lower=1> J;                  // number of historical oHCM trials (4)
  vector[J] y;                     // placebo-adjusted KCCQ effects
  vector<lower=0>[J] sigma;        // SEs after discounting: se / sqrt(w)

  real mu0;                        // prior mean of mu            (0)
  real<lower=0> sigma_mu;          // prior SD of mu              (5)
  real<lower=0> tau0;              // HalfNormal scale for tau    (1.5)

  real<lower=0> beta_sd;           // attenuation scale, oHCM -> nHCM (2)
  real<lower=0> base_rate_alpha;   // Beta prior on Pr(transfer)  (3.5)
  real<lower=0> base_rate_beta;    //                             (1.5)
  real<lower=0> sigma_null;        // SD of the null component    (2)
}
```

Stan: the hierarchical part

```
parameters {
  real mu;                                  // oHCM class mean
  real<lower=1e-6> tau;                     // between-trial SD
  vector<offset=mu, multiplier=tau>[J] theta;   // non-centered
}
model {
  mu  ~ normal(mu0, sigma_mu);                  // N(0, 5)
  tau ~ normal(0, tau0);                        // HalfNormal(1.5)
  theta ~ normal(mu, tau);
  y ~ normal(theta, sigma);                     // sigma = se / sqrt(w)
}
```

Stan: the transfer mixture

```
generated quantities {
  real theta_oHCM = normal_rng(mu, tau);            // new oHCM trial
  real pi_active = beta_rng(base_rate_alpha, base_rate_beta);  // prior only
  real<lower=0> beta_atten = abs(normal_rng(0, beta_sd));   // prior only
  int<lower=0, upper=1> Z = bernoulli_rng(pi_active);       // prior only
  real theta_nHCM = (Z == 1) ? theta_oHCM - beta_atten
                             : normal_rng(0, sigma_null);
  real lprior = normal_lpdf(mu | mu0, sigma_mu) + normal_lpdf(tau | 0, tau0);
  // ... plus y_rep (PPC) and log_lik (LOO)
}
```

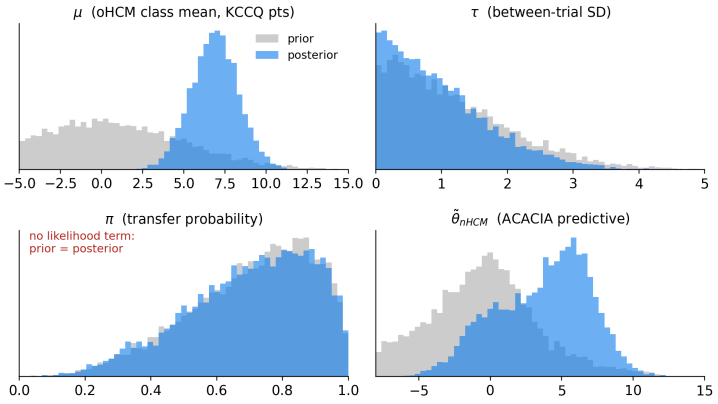
Python: fit, gate, freeze

```
fit, verdict = fit_stan(STAN_FILE, stan_data, seed=12345, adapt_delta=0.98)
# 0 divergences, R-hat <= 1.01, bulk/tail ESS >= 400 -- or it raises

theta = fit.stan_variable("theta_nHCM")
kccq = build_endpoint_posterior(
    "kccq_css", "KCCQ CSS points", draws=theta, mid_threshold=5.0,
    beneficial_direction="greater")

freeze(prediction_id="pred:acacia-hcm:20260422:e8a10488", trial=ACACIA_SPEC,
       endpoints={"kccq_css": kccq, "peak_vo2": pvo2},
       probability_of_success={"overall": 0.572, "success_rule": "either"},
       evidence_cutoff_date="2026-04-22", model_version="mixture v0.1", ...)
# append-only; stamps git SHA + SHA-256 of the Stan and data files
```

How much did we learn?



μ moves from 0 to 6.9 (4.0–9.7); τ shrinks; π does not move at all.

Diagnostics and a three-fold backtest

Sampler checks

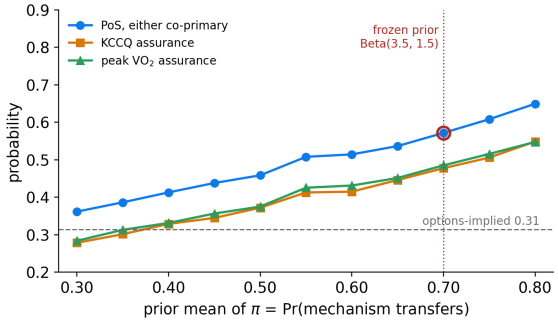
check	value	gate
divergences	0	0
max $\hat{R}$	1.003	1.01
min bulk ESS	2923	400
min tail ESS	2888	400

Leave-one-trial-out (oHCM, same model)

held out	PoS	observed	pred $\rightarrow$ obs
EXPLORER	0.76	success	7.0 $\rightarrow$ 7.6
VALOR	0.39	success	9.4 $\rightarrow$ 7.0
SEQUOIA	0.65	success	7.0 $\rightarrow$ 7.9

4 chains $\times$ 2000 draws; PPC: all 4 trials inside the 95% replicate interval. Backtest: Brier 0.18, log score 0.54, 2/3 inside the 95% predictive — a pipeline check, not calibration.

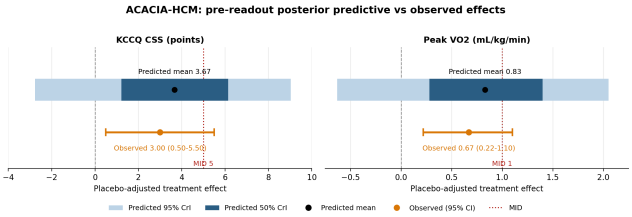
Sensitivity



All four borrowing weights scaled

	0.75×	**1.00×**	1.25×
μ posterior mean	6.70	**6.85**	6.95
KCCQ predictive mean	3.60	**3.67**	3.78
Pr(KCCQ $>$ MID)	0.37	**0.40**	0.42
PoS, either hits	0.57	**0.57**	0.58

Readout: scored



	KCCQ	VO_2
predicted mean	3.67	0.83
observed	3.0	0.67
quantile rank	0.40	0.40
z	−0.21	−0.22

Right call (PoS 0.57, success: Brier 0.18, log score 0.56); effects ~22% optimistic; predictive 2–3× wider than the realized CI.

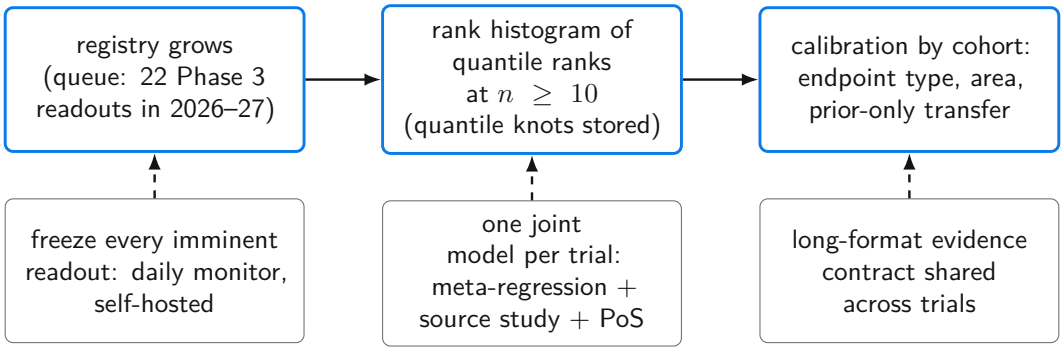
Market-implied probabilities: a different question

$$f_{\mathbb{Q}}(S_T) = e^{rT}\,\frac{\partial^2 C(K,T)}{\partial K^2}\Big|_{K=S_T}, \qquad \Pr_{\mathbb{Q}}(S_T > S^*) = \int_{S^*}^{\infty} f_{\mathbb{Q}}(S_T)\,dS_T, \qquad S^* = 1.10\times S_0$$

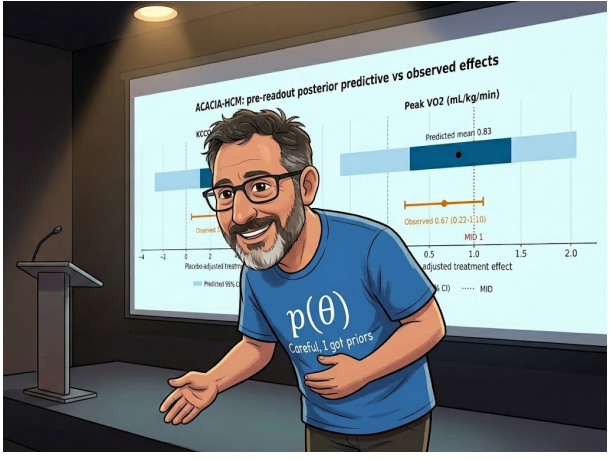
CYTK, options snapshot 2026-04-21 (spot \$65.80)	Pr(close > \$72.38)	80% band
June 18 expiry	0.12	0.04–0.19
July 17 expiry	0.32	0.17–0.47
clinical PoS, either co-primary hits	0.57	
realized 2026-05-05 close	**\$77.76, +18%**	

Risk-neutral, not clinical. The option price carries risk premia, commercial value and timing; it answers "will the stock clear a threshold by expiry", not "will the trial hit".

What's next


registry grows
(queue: 22 Phase 3
readouts in 2026–27)
rank histogram of
quantile ranks
at $n \geq 10$
(quantile knots stored)
calibration by cohort:
endpoint type, area,
prior-only transfer
freeze every imminent
readout: daily monitor,
self-hosted
one joint
model per trial:
meta-regression +
source study + PoS
long-format evidence
contract shared
across trials

Thank you



Eric Novik

Generable · NYU

eric.novik@nyu.edu

linkedin.com/in/enovik